

SEQUENCE LISTING

<110> INVITROGEN CORPORATION
DALBY, Brian
BENNETT, Robert

<120> DELIVERY OF FUNCTIONAL PROTEIN SEQUENCES
BY TRANSLOCATING POLYPEPTIDES

<130> INVIT1280-1

<150> PCT/US00/08571

<151> 2000-03-31

<150> 60/127,467

<151> 1999-03-31

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 6404

<212> DNA

<213> Artificial Sequence

<220>

<223> vector pVP22/Myc-His

<400> 1

gacggatcgg	gagatctccc	gateccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttggtgtgt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttaggggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacgggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccgcg	tggctgaccg	cccaacgacc	360
cccgcacatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgcccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaattggccc	gcctggcatt	540
atgcccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaattgggcg	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagt	900
taagcttatt	atgacctctc	gccgctccgt	gaagtcgggt	ccgcgggagg	ttccgcgcga	960
tgagtacgag	gatctgtact	acaccccgtc	ttcaggatat	gcgagtcctg	atagtcgcgc	1020
tgacacctcc	cgccgtggcg	ccctacagac	acgctcgcgc	cagagggggc	aggtccggtt	1080
cgtccagtac	gacgagtcgg	attatgccct	ctacgggggc	tcgtcttccg	aagacgacga	1140
acacccggag	gtcccccgga	cgcggcgctc	cggtttccgg	gcgggtttgt	ccggcccggg	1200
gcctgcgcgg	gcgcctccgc	cacccgctgg	gtccggaggg	gccggacgca	cacccaccac	1260
cgcccccccg	gccccccgaa	cccagcgggt	ggcgtctaag	gccccgcg	ccccggcg	1320
ggagaccacc	cgcggcagga	aatcgggcca	gccagaatcc	gccgcactcc	cagacgcccc	1380
cgcgctcgag	gcgccaaccc	gatccaagac	acccgcgcag	gggctggcca	gaaagctgca	1440
ctttagcacc	gcccccccaa	accccgacgc	gcatgggacc	ccccgggtgg	ccggctttaa	1500
caagcgcgtc	ttctgcgcgc	cggtcggggc	cctggcgccc	atgcatgccc	ggatggcg	1560
tgctcagctc	tgggacatgt	cgcgtccg	cacagacgaa	gacctcaacg	aactccttgg	1620
catcaccacc	atccgcgtga	cggctctgca	gggcaaaaac	ctgcttcagc	gcgccaacga	1680
gttggtgaat	ccagacgtgg	tgcaggacgt	cgacgcggcc	acggcgactc	gagggcgctc	1740

202070-4E84E660

tgcggcgctcg	cgccccacccg	agcgacctcg	agccccagcc	cgctccgctt	ctcgccccag	1800
acggcccgctc	gaggggtaccg	agctcggatc	cactagtcca	gtgtgggtgga	attctgcaga	1860
tatccagcac	agtggcgggc	gctcgagtct	agagggcccc	cggttcgaac	aaaaactcat	1920
ctcagaagag	gatctgaata	tgcataccgg	tcatcatcac	catcaccatt	gagtttaaac	1980
ccgctgatca	gcctcgactg	tgccttctag	ttgccagcca	tctgttggtt	gccccctccc	2040
cgtgccttcc	ttgaccctgg	aaggtgccac	tcccactgtc	ctttcctaata	aaaatgagga	2100
aattgcatcg	cattgtctga	gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	2160
cagcaagggg	gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	2220
ggcttctgag	gcggaaagaa	ccagctgggg	ctctaggggg	tatccccacg	cgccctgtag	2280
cggcgcattha	agcgcgggcg	gtgtgggtgg	tacgcgcagc	gtgaccgcta	cacttgccag	2340
cgccctagcg	cccgcctcct	tcgctttctt	cccttccttt	ctcgccacgt	tcgcccggctt	2400
tccccgtcaa	gctctaaatc	ggggcatccc	tttaggggtc	cgatttagtg	ctttacggca	2460
cctcgacccc	aaaaaacttg	attaggggtg	tggttcacgt	agtgggccat	cgccctgata	2520
gacggttttt	cgccctttga	cgttggagtc	cacgttcttt	aatagtggac	tcttggtcca	2580
aactggaaca	acactcaacc	ctatctcggt	ctattctttt	gatttataag	ggattttggg	2640
gatttcggcc	tattgggttaa	aaaatgagct	gatttaacaa	aaatttaacg	cgaatttaatt	2700
ctgtggaatg	tgtgtcagtt	aggggtgtgga	aagtccccag	gctccccagg	caggcagaag	2760
tatgcaaagc	atgcatctca	attagtcagc	aaccagggtg	ggaaagtccc	caggctcccc	2820
agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccatag	tcccgcacct	2880
aactccgccc	atcccgcgcc	taactccgcc	cagttccgce	cattctccgc	cccatggctg	2940
actaattttt	tttattttatg	cagagggcga	ggcgcctct	gcctctgagc	tattccagaa	3000
gtagtggaga	ggcttttttg	gagggcctag	cttttgcaaa	aagctcccgg	gagcttgtat	3060
atccattttc	ggatctgata	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	3120
tggattgcac	gcaggttctc	cggccgcttg	ggtggagagg	ctattcggct	atgactgggc	3180
acaacagaca	atcggctgct	ctgatgccgc	cgtgttccgg	ctgtcagcgc	aggggcgccc	3240
ggttcttttt	gtcaagaccg	acctgtccgg	tgccctgaat	gaactgcagg	acgaggcagc	3300
gcggctatcg	tggctggcca	cgacggggcg	tccttgccga	gctgtgctcg	acgttgtcac	3360
tgaagcggga	agggactggc	tgctattggg	cgaagtgccg	gggcaggatc	tcctgtcatc	3420
tcaccttgct	cctgcccaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	3480
gcttgatcg	gctacctgcc	cattcgacca	ccaagcgaag	catcgcatcg	agcgagcacg	3540
tactcggatg	gaagccggtc	tttcgcatca	ggatgattcg	gacgaagagc	atcaggggct	3600
cgcgccagcc	gaactgttgc	ccaggctcaa	ggcgcgcag	ccgcacggcg	aggatctcgt	3660
cgtgacccat	ggcgatgcct	gcttgccgaa	tatcatgggtg	gaaaatggcc	gcttttctgg	3720
attcatcgac	tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cgttggtctac	3780
ccgtgatatt	gctgaagagc	ttggcggcga	atgggctgac	cgcttcctcg	tgctttacgg	3840
tatcgccgct	cccgatccgc	agcgcacatc	cttctatcgc	cttcttgacg	agttcttctg	3900
agcgggactc	tgggggttcgc	gaaatgaccg	accaagcgac	gcccacacctg	ccatcacgag	3960
atttcgattc	caccgcgcgc	ttctatgaaa	ggttgggctt	cggaatcggt	ttccgggacg	4020
cggctggatg	gatcctccag	cgcggggatc	tcatgctgga	gttcttcgcc	caccccaact	4080
tgtttattgc	agctttataat	ggttacaaat	aaagcaatag	catcacaaat	ttcacaaata	4140
aagcattttt	ttcactgcac	tctagttgtg	gtttgtccaa	actcatcaat	gtatcttata	4200
atgtctgtat	accgtcgacc	tctagctaga	gcttggcgta	atcatgggtca	tagctgtttc	4260
ctgtgtgaaa	ttgttatccg	ctcacaaatc	cacacaacat	acgagccgga	agcataaagt	4320
gtaaagcctg	gggtgcctaa	tgagtgaagt	aactcacatt	aattgcgttg	cgctcactgc	4380
ccgctttcca	gtcgggaaac	ctgtcgtgcc	agctgcatta	atgaatcggc	caacgcgcgg	4440
ggagaggcgg	tttgcgatatt	gggcgctctt	ccgcttcctc	gctcactgac	tcgctgcgct	4500
cggctcgttcg	gctgcggcga	gcgggtatcag	ctcactcaaa	ggcggttaata	cgggttatcca	4560
cagaatcagg	ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	4620
accgtaaaaa	ggccgcggtg	ctggcgtttt	tccataggtc	ccgccccctt	gacgagcatc	4680
acaaaaatcg	acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	agataaccagg	4740
cgtttccccc	tggaaagctcc	ctcgtgcgct	ctcctgttcc	gaccctgccg	cttaccggat	4800
acctgtccgc	ctttctccct	tcgggaagcg	tggcgctttc	tcaatgctca	cgctgtaggt	4860
atctcagttc	ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccggttc	4920
agcccgaccg	ctgcgcctta	tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	4980
acttatcgcc	actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	5040
gtgctacaga	gttcttgaag	tgggtggccta	actacggcta	cactagaagg	acagtatttg	5100
gtatctgcgc	tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	5160
gcaaacacaa	caccgctggg	agcggtggtt	tttttgggtg	caagcagcag	attacgcgca	5220
gaaaaaaagg	atctcaagaa	gatcctttga	tcttttctac	ggggtctgac	gctcagtgga	5280
acgaaaactc	acgttaaggg	attttgggtc	tgagattatc	aaaaaggatc	ttcacctaga	5340
tccttttaaa	ttaaaaatga	agttttaaat	caatctaaag	tatatatgag	taaacttggt	5400

ctgacagtta	ccaatgctta	atcagtgagg	cacctatctc	agcgatctgt	ctatttcggt	5460
catccatagt	tgccctgactc	cccgtcgtgt	agataactac	gatacgggag	ggcttaccat	5520
ctggccccag	tgctgcaatg	ataccgcgag	acccacgctc	accggctcca	gatttatcag	5580
caataaacca	gccagccgga	agggccgagc	gcagaagtgg	tcctgcaact	ttatccgcct	5640
ccatccagtc	tattaattgt	tgccgggaag	ctagagtaag	tagttcgcca	gttaatagtt	5700
tgcgcaacgt	tggtgccatt	gctacaggca	tcgtgggtgtc	acgctcgtcg	tttggtatgg	5760
cttcattcag	ctccgggttc	caacgatcaa	ggcgagttac	atgatcccc	atgttgtgca	5820
aaaaagcgg	tagctccttc	ggtcctccga	tcgttgtcag	aagtaagttg	gccgcagtgt	5880
tatcactcat	ggttatggca	gcactgcata	attctcttac	tgtcatgcc	tccgtaagat	5940
gcttttctgt	gactggtgag	tactcaacca	agtcattctg	agaatagtgt	atgcgggcgac	6000
cgagttgctc	ttgcccggcg	tcaatacggg	ataataccgc	gccacatagc	agaacttta	6060
aagtgtctac	cattggaaaa	cgctcttcgg	ggcgaaaact	ctcaaggatc	ttaccgctgt	6120
tgagatccag	ttcgatgtaa	cccactcgtg	cacccaactg	atcttcagca	tcttttactt	6180
tcaccagcgt	ttctgggtga	gcaaaaaacg	gaaggcaaaa	tgccgcaaaa	aagggaataa	6240
gggcgacacg	gaaatgttga	atactcatat	tcttctcttt	tcaatattat	tgaagcattt	6300
atcagggtta	ttgtctcatg	agcggatata	tatttgatg	tatttagaaa	aataaataaa	6360
taggggttcc	gcgcacattt	ccccgaaaag	tgccacctga	cgtc		6404

<210> 2

<211> 6420

<212> DNA

<213> Artificial Sequence

<220>

<223> vector pVP22/Myc-His-TOPO

<400> 2

gacggatcgg	gagatctccc	gatccccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctcccctg	cttgtgtgtt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacgggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggccccgcc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgcca	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaattggccc	gcctggcatt	540
atgcccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtatttagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcgggttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaattggcg	780
gtaggcgtgt	acgggtgggag	gtctatatata	gcagagctct	ctggctaact	agagaacca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagt	900
taagcttatt	atgacctctc	gccgctccgt	gaagtcgggt	ccgcgggagg	ttccgcgcga	960
tgagtacgag	gatctgtact	acaccccgtc	ttcaggtatg	gcgagtcgcc	atagtcgcc	1020
tgacacctcc	cgccgtggcg	ccctacagac	acgctcgcgc	cagaggggag	aggtccggtt	1080
cgtccagtac	gacgagtcgg	attatgccct	ctacggggggc	tcgtcttcgc	aagacgacga	1140
acacccggag	gtcccccgga	cgcggcgctc	cgtttccggg	gcggttttgt	ccggccccgg	1200
gctgcgcg	gcgcctccgc	cacccgctgg	gtccggaggg	gccggacgca	cacccaccac	1260
cgcccccg	gcccccgaa	cccagcgggt	ggcgactaag	gcccccgcg	ccccggcg	1320
ggagaccacc	cgcggcagga	aatcggccca	gccagaatcc	gccgcactcc	cagacgcccc	1380
cgcgtcgacg	gcgccaaccc	gatccaagac	acccgcgcag	gggctggcca	gaaagctgca	1440
ctttagcacc	gccccccaa	accccgacgc	gccatggacc	ccccgggtgg	ccggcttta	1500
caagcgcgtc	ttctgcgccg	cggtcggggc	cctggcgggc	atgcatgccc	ggatggcggc	1560
ggtccagctc	tgggacatgt	cgcgtccgcg	cacagacgaa	gacctcaacg	aactccttgg	1620
catcaccacc	atccgcgtga	cggctctgca	gggcaaaaac	ctgcttcagc	gcgccaacga	1680
gttggtgaat	ccagacgtgg	tgcaggacgt	cgacgcggcc	acggcgactc	gagggcggtt	1740
tgcggcgctc	cgccccaccg	agcgacctgc	agccccagcc	cgctccgctt	ctcgccccag	1800
acggcccgtc	gagggatacc	agctcgagtc	cactagtcca	gtgtggtgga	attgcctta	1860
agggaatttc	tgcgatatcc	cagcagagtg	gcggccgctc	gagtcctagag	ggccccgggt	1920
tcgaacaaaa	actcatctca	gaagaggatc	tgaatatgca	taccggctcat	catcaccatc	1980
accattgagt	ttaaaccgcg	tgatcagcct	cgactgtgcc	ttctagttgc	cagccatctg	2040

ttgtttgccc	ctcccccg	ccttccctga	ccctggaagg	tgccactccc	actgtccttt	2100
cctaataaaa	tgaggaaatt	gcacgcgatt	gtctgagtag	gtgtcattct	attctggggg	2160
gtgggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	2220
atgcggtggg	ctctatggct	tctgaggcgg	aaagaaccag	ctggggctct	aggggggtatc	2280
cccacgcgcc	ctgtagcgcc	gcattaagcg	cggcgggtgt	ggtggttacg	cgcagcgatga	2340
ccgctacact	tgccagcgcc	ctagcgcccg	ctcctttcgc	tttcttccct	tcctttctcg	2400
ccacgttcgc	cggctttccc	cgtcaagctc	taaatcgggg	catcccttta	gggttccgat	2460
ttagtgtttt	acggcacctc	gaccccaaaa	aacttgatta	gggtgatggg	tcacgtagtg	2520
ggccatcgcc	ctgatagacg	gtttttcgcc	ctttgacgtt	ggagtccacg	ttctttaata	2580
gtggactctt	gttcctaaact	ggaacaacac	tcaaccctat	ctcggctctat	tcttttgatt	2640
tataagggat	tttggggatt	tgggcctatt	ggttaaaaaa	tgagctgatt	taacaaaaat	2700
ttaacgcgaa	ttaattctgt	ggaatgtgtg	tcagttaggg	tgtggaaagt	ccccaggctc	2760
cccaggcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	aggtgtggaa	2820
agtcctccagg	ctccccagca	ggcagaagta	tgcaaaagcat	gcatctcaat	tagtcagcaa	2880
ccatagtccc	gcccctaact	cgccccatcc	cgccccctaac	tccgcccagt	tccgcccatt	2940
ctccgccccca	tggctgacta	atTTTTTTT	tttatgcaga	ggccgaggcc	gcctctgcct	3000
ctgagctatt	ccagaagtag	tgaggaggct	tttttgagg	cctaggcttt	tgcaaaaagc	3060
tcccgggagc	ttgtatatcc	atTTTtcggat	ctgatcaaga	gacaggatga	ggatcgtttc	3120
gcatgattga	acaagatgga	ttgcacgcag	gttctccggc	cgcttgggtg	gagaggctat	3180
tcggctatga	ctgggcacaa	cagacaatcg	gctgctctga	tgccgccgtg	ttccggctgt	3240
cagcgcaggg	gcgcccgggt	ctttttgtca	agaccgacct	gtccgggtgcc	ctgaatgaac	3300
tgaggagcgg	ggcagcgcg	ctatcgtggc	tggccacgac	gggcgttcc	tgccgagctg	3360
tgtctgacgt	gtctactgaa	cggggaagg	actggtctgct	attgggcgaa	gtgcccgggc	3420
aggatctcct	gtcatctcac	cttgtctctg	ccgagaaagt	atccatcatg	gctgatgcaa	3480
tgccggcggt	gcatacgctt	gatccggcta	cctgcccatt	cgaccaccaa	gcgaaacatc	3540
gcatcgagcg	agcacgtact	cggatggaag	ccggtcttgt	cgatcaggat	gatctggacg	3600
aagagcatca	ggggctcgcg	ccagccgaac	tgttcgccag	gctcaaggcg	cgcagccccg	3660
acggcgagga	tctcgctcg	acccatggcg	atgcctgctt	gccgaatatc	atggtggaaa	3720
atggccgctt	ttctggattc	atcgactgtg	gccggctggg	tgtggcggac	cgctatcagg	3780
acatagcggt	ggctaccgct	gatattgctg	aagagcttgg	cggcgaaatg	gctgaccgct	3840
tcctcgctgt	ttacgggtatc	ggcgtcccc	attcgagcg	catcgccctc	tatcgccctc	3900
ttgacgagtt	cttctgagcg	ggactctggg	gttcgcgaaa	tgaccgacca	agcgacgccc	3960
aacctgccat	cacgagattt	cgattccacc	gccgccttct	atgaaagggt	gggcttcgga	4020
atcgtttttc	gggacgcgg	ctggatgatc	ctccagcgcg	gggatctcat	gctggagttc	4080
ttcgccacc	ccaacttggt	tattgcagct	tataatgggt	acaaataaag	caatagcatc	4140
acaaatttca	caaataaagc	atTTTTTTT	ctgcattcta	gttgtgggtt	gtccaaactc	4200
atcaatgtat	cttatcatgt	ctgtataacc	tcgacctcta	gctagagctt	ggcgtaatca	4260
tggtcatagc	tgtttctctg	gtgaaattgt	tatccgctca	caattccaca	caacatacga	4320
gccggaagca	taaagtgtaa	agcctggggg	gcctaattgag	tgagctaact	cacattaatt	4380
gcgttgccgt	cactgcccgc	ttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	4440
atcgcccaac	gcgcggggag	aggcggtttg	cgtattgggc	gctcttccgc	ttcctcgctc	4500
actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	4560
gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	4620
cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	4680
ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccagacagga	4740
ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	4800
ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcggtggc	gctttctcaa	4860
tgctcacgct	gtaggtatct	cagttcggtg	taggtcgctc	gctccaagct	gggctgtgtg	4920
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtc	4980
aaccgggtaa	gacacgactt	atcgccactg	gcagcgacca	ctggtaacag	gattagcaga	5040
gcgaggatag	taggcgggtg	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	5100
agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	5160
ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtgggttttt	tggttgcaag	5220
cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	5280
tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	5340
aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	ttaaatcaat	ctaaagtata	5400
tatgagtaaa	cttgggtctga	cagttacca	tgcttaatca	gtgaggcacc	tatctcagcg	5460
atctgtctat	ttcggttcac	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	5520
cgggagggt	taccatctgg	ccccagtgct	gcaatgatac	cgcgagaccc	acgctcaccc	5580
gctccagatt	tatcagcaat	aaaccagcca	gccggaagg	ccgagcgag	aagtggctct	5640
gcaactttat	ccgcctccat	ccagtctatt	aattggtgcc	gggaagctag	agtaagtagt	5700

```

tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc 5760
tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 5820
tcccccatgt tgtgcaaaaa agcggtttagc tccttcggtc ctccgatcgt tgtcagaagt 5880
aagttggccg cagtgttata actcatgggt atggcagcac tgcataattc tcttactgtc 5940
atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 6000
tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtaa tacgggataa taccgcgcca 6060
catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 6120
aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc caactgatct 6180
tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 6240
gcaaaaaagg gaataagggc gacacggaaa tggtgaatac tcatactctt cctttttcaa 6300
tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6360
tagaaaaata aacaaatagg ggttccgcgc acatttccc gaaaagtgcc acctgacgtc 6420

```

```

<210> 3
<211> 12
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Peptide linker

```

```

<400> 3
Leu Ala Arg Leu Leu Ala Arg Leu Leu Ala Arg Leu
  1             5             10

```

```

<210> 4
<211> 67
<212> PRT
<213> Artificial sequence

```

```

<220>
<223> Conserved sequence of steroid/thyroid hormone
      receptor superfamily DNA-binding domain

```

```

<221> VARIANT
<222> (0)...(0)
<223> Xaa is non-conserved amino acids within the
      DNA-binding domain

```

```

<221> VARIANT
<222> 7, 9, 11, 13, 22, 27, 58, 61, 66
<223> amino acid residues that are almost universally
      conserved, but for which variations have been
      found in some identified hormone receptors

```

```

<400> 4
Cys Xaa Xaa Cys Xaa Xaa Asp Xaa Ala Xaa Gly Xaa Tyr Xaa Xaa Xaa
  1             5             10             15
Xaa Cys Xaa Xaa Cys Lys Xaa Phe Phe Xaa Arg Xaa Xaa Xaa Xaa Xaa
      20             25             30
Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Lys
      35             40             45
Xaa Xaa Arg Xaa Xaa Cys Xaa Xaa Cys Arg Xaa Xaa Lys Cys Xaa Xaa
      50             55             60
Xaa Gly Met
65

```

```

<210> 5
<211> 34
<212> DNA

```

202070-LEBZ660

<400> 11
 Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser
 1 5 10

<210> 12
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 12
 Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly
 1 5 10

<210> 13
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 13
 Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Ser Gly Ser Thr
 1 5 10 15
 Lys Gly

<210> 14
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 14
 Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly
 1 5 10

<210> 15
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 15
 Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Thr
 1 5 10 15
 Lys Gly

<210> 16
 <211> 14
 <212> PRT
 <213> Artificial Sequence

2020/07/07 14:00:00

<220>

<223> Peptide linker moiety

<400> 16

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Glu Phe
1 5 10

<210> 17

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<400> 17

Ser Arg Ser Ser Gly
1 5

<210> 18

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<400> 18

Ser Gly Ser Ser Cys
1 5

<210> 19

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<400> 19

Met Gly Arg Ser Gly Gly Gly Cys Ala Gly Asn Arg Val Gly Ser Ser
1 5 10 15
Leu Ser Cys Gly Gly Leu Asn Leu Gln Ala Met
20 25

<210> 20

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<221> VARIANT

<222> (0)...(0)

<223> Xaa is (GmS)n, where m is
from 2 to 4 and n is from 1 to 11.

<400> 20

Ala Met Xaa Ala Met

00937837 010702

1

5

<210> 21

<211> 16

<212> PRT

<213> Drosophila acanthoptera

<400> 21

Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys
1 5 10 15

09937837.010702